# User and service equilibrium in a structural model of traffic assignment to a transit network

Fabien Leurent, Ektoras Chandakas, Alexis Poulhès

14th EWGT Conference, Poznan, 6-9 September 2011

# User and service equilibrium in a structural model of traffic assignment to a transit network

Fabien Leurent, Ektoras Chandakas, Alexis Poulhès

*Université Paris-Est, Laboratoire Ville Mobilité Transport, Ecole des Ponts ParisTech*
*6-8 avenue Blaise Pascal, 77455 Marne la Vallée Cédex 2, France*

**Abstract**

A structural model is provided for traffic assignment to a transit model. It deals with a hierarchy of layers from passenger to network passing by vehicle, service route and line, while taking advantage of the spatial structure of service routes and lines. A range of capacity effects are addressed: in-vehicle passenger capacity, access-egress capacity in relation to dwelling time, platform occupancy by vehicles which may reduce service frequency.
The model treatment involves a line sub-model, which amounts to a sophisticated cost-flow relationship, and platform sub-models first for passenger storage and flowing, second for vehicle operations. Traffic equilibrium is addressed in a hyperpath framework on the basis of leg links.


*Keywords: Transit capacity; Line model; Platform model; Dwell time; Frequency reduction; Zip algorithm*

## 1. Introduction

In the transit network of a large urban area, it frequently occurs that a high duty line is submitted to heavy congestion at the peak hours on working days, especially so at the morning peak in the central part of the urban area. Under that circumstance, not only may the passengers experience the discomfort of crowding, delay and unreliability, but also the operation of services may be disrupted by increased dwelling times, vehicle bunching and delays, leading to reductions in service frequency. Thus transit traffic is inconvenienced and disrupted at both levels of mobile units, passenger and vehicle.

Although many transit engineers are well aware of these issues, which are discussed at length in the Transit Quality of Service and Capacity Manual (Trb, 2003), little attention has been given by the scientific community of transportation modellers to the interplay of passenger and vehicle traffic. On one hand, passenger traffic is addressed by models of traffic assignment to a transit network, typically in planning studies where the operating conditions are described through the service routes and operation frequencies, section run time, station dwelling time, maybe vehicle capacity (Thomas, 1991; Ortuzar and Willumsen, 2004); in some advanced models, one or several of these features are taken as flow-responsive. However little consideration has been given to the dynamic conditions of real-time operations: the flow-responsive features and policies are described mostly by heuristic constructs rather than physical and technical models. Route frequency has been related by Lam et al (1999) to fleet size and cycle time by vehicle, taking both station dwelling time and section running time as flow dependent. In fact however, the scarce resource pertains to platform availability rather than fleet size. Cepeda et al (2006) associated to each station along a service route a fictive frequency adapted from nominal to account for the effects of the incoming flow and the residual capacity: but the relationship is an artefact, with neither account of the egress flow on the dwelling time nor forward propagation of the reduction in service frequency. The macroscopic dynamic assignment models proposed so far (eg. Sumalee et al, 2009) take the service timetable as exogenous, in contradiction with the unreliability that appears under severe congestion. Micro-simulation is still restricted to local issues of station design (eg. Hoogendorn and Daamen, 2005).

On the other hand, service operations are planned at the line level on the basis of a model of vehicle running along the service trajectory, taking into account the local conditions such as line geometry and station dwelling time. The passenger traffic is considered only through its influence on the dwelling times and the vehicle load (hence mass) (e.g. Vuchic, 2006; Lai et al, 2011).

This paper is purported to bridge the gap between these two families of models. It brings out a model of traffic assignment to a transit network that is sensitive to the interplay between passenger and vehicle traffic at the line level. Platform occupancy by successive vehicles influences service frequency. Dwelling time depends on passenger access and egress flows. A local decrease in frequency implies a decrease in service residual capacity that is delivered downstream, all the more so as it comes along with a decrease in the residual capacity in each vehicle. On the passenger side, local queuing may occur on station platform and contribute to leg cost, the leg being the trip part along the transit line from access to egress station. Route choice at the network level depends on leg costs.

Our approach to the physical and economic features is static, in order to deal with service frequency conveniently. Transit modes such as bus or train are distinguished in their travel conditions, including platform layout and the potential location of route choice within a line station. Route services are grouped into a line of operation inasmuch as they share resources of section track and station platform, which makes their operation interdependent. The tree-like pattern of a line provides our main line of attack, as in Leurent's model of seat capacity (2006, 2010): from upstream to downstream the passenger flow constraints are progressively tied up along the successive stations, with influence on service frequency and residual capacity. Conversely to this 'zip' treatment an 'unzip' treatment from downstream to upstream yields the travel conditions by trip leg, thus providing the basis for passenger route choice by origin-destination pair at the network level.

The rest of the paper is organized into five sections. First, the model assumptions are set out in Section 2. Then, Section 3 provides the line sub-model, including the 'zip' and 'unzip' algorithms. The model composition and overall solution scheme is sketched out in Section 4. Section 5 deals with a classroom case adapted from the busiest railway line in the Paris metropolitan area (line A of the RER, the so-called Regional Express Network). Lastly, Section 6 concludes about the model contents and points to potential developments.

## 2. Model assumptions

As stated by Leurent (2011a) a transit system involves five main subsystems: Line, Vehicle, Station, Passenger and Management. Each one deserves special consideration in the modelling.

### *2.1. Network topology, from infrastructure to services*

The *infrastructure network* is made up of a set $N$ of nodes $n$, together with a set A of arcs $a$ (unidirectional links) with endpoints in $N$. Basically, an arc represents a state transition for a trip-maker: set $A$ is split between subsets $A_B$ of boarding links, $A_A$ of alighting links, $A_I$ of interstation links and $A_S$ of sojourn links within vehicle and station and $A_P$ of other private links for pedestrian access or transfer between two services routes. Such a route is modelled as an oriented sequence of sojourn and interstation links, alternately.

In the basic hyperpath model of transit assignment (Spiess and Florian, 1989), a passenger routing structure is in fact a bundle of paths from origin to destination: the structure may combine service routes at given choice nodes, i.e. platform nodes (or information nodes) at which point several routes are available and attractive to get the trip-maker closer to his destination. This is a weak interaction between service routes, as compared with strong interaction on the operator side where given physical nodes and arcs – notably sojourn and interstation – are shared by several services. Here it is assumed that such related services belong to the same "service line", i.e. a sub-network that connects platform user nodes from entry to exit.

Thus the trip-maker is faced to a *service network* with one leg arc per couple of entry-exit stations along any service route. Such a leg arc represents one or several interstation and sojourn links along the service route in the infrastructure network that is still considered on the supply side. On the demand side, it is assumed that the passengers consider only leg arcs and private arcs for route boarding or route alighting, station access and transfer from route to route at a given interchange station (e.g. De Cea and Fernandez, 1993); Leurent, 2011a).

Overall, the *service network* is made up of the set $A_S \equiv A_P \bigcup A_B \bigcup A_A \bigcup A_L$, $A_L$ being the set of service legs, with the associated node sets.

## 2.2. Vehicle features

By service route $z$, the average vehicle has an on-board capacity of $k_z$ passengers. At station $n$, the time for a passenger to pass a vehicle door is $t_{zn}^{-}$ on egress and $t_{zn}^{+}$ on access. Denoting by $y_{za}$ the number of passengers on-board on link $a$ and by $y_{zn}^{-}$ (resp. $y_{zn}^{+}$) the number of passengers alighting (resp. boarding) at station $n$, then the time for passenger discharge and refill cannot be less than

$$T_{zn}^{-+} \equiv (y_{zn}^{-}.t_{zn}^{-} + y_{zn}^{+}.t_{zn}^{+}) / X_z ,$$

$X_z$ being the number of passengers that can simultaneously traverse the vehicle door(s). Let us define $g_{zn}^{-} \equiv t_{zn}^{-} / X_z$ and $g_{zn}^{+} \equiv t_{zn}^{+} / X_z$. The $k_z$ and $X_z$ parameters depend on the vehicle design, whereas $t_{zn}^{-}$ and $t_{zn}^{+}$ also involve the station design.

## 2.3. Platform issues: the vehicle side

The station platform $n$ of a given line, $\ell$, is a resource to be shared between the vehicles of the service routes $z \in \ell$. During a period of given time length, say $H$, let $f_{zn}$ denote the frequency of operation of route $z$ at node $n$, $T_{zn} \equiv \max\{T_{zn}^{1} + T_{zn}^{-+}, T_{zn}^{0}\}$ be the dwell time of a vehicle (denoting by $T_{zi}^{1}$ a dead time for door operation etc) and $\omega_{zn}$ be the minimum inter-vehicular gap prior to a vehicle servicing $z$ (on average). The resource occupancy must satisfy the following constraint:

$$\sum_{z \in \ell \cap n} f_{zn} (T_{zn} + \omega_{zn}) \leq H \ .$$

The minimum inter-vehicular gap is set up by the operator on the basis of the station and its upstream inter-station, their geometry, track features, signaling system and so on.

## 2.4. Line traffic

A tree-like structure is assumed for each line $\ell$, with branches connecting on a main trunk providing service between pairs $(i, j)$ of stations that are oriented along the line: this is denoted as $i >_{\ell} j$ ($i$ is upstream of $j$). The set of stations serviced by route $z$ between $i$ and $j$ is denoted as $z \bigcap [i, j[ \, : \, i, j \in S_{\ell}$.

At the vehicle level, let $t_{za}$ denote the running time along section $a$ for route $z$.

Then the vehicle time by route $z$ from $i$ to $j$ is

$$V_{ij}^{z} = \sum_{n \in z \cap [i,j[} T_{zn} + \sum_{a \in z \cap [i,j[} t_{za} \ .$$

Notation $a \in z \bigcap [i, j[$ is an unambiguous shorthand to encompass the sections travelled along $z$ from $i$ to $j$.

The traffic of passengers along the line is analyzed at two levels external vs. internal. The internal perspective involves the line sub-model that is addressed in the next Section. On the external perspective, during a given time period the line carries a vector $\mathbf{q}^{\ell} = [q_{ij}^{\ell} : i >_{\ell} j]$ of passenger flows by leg and yields leg costs $\mathbf{c}^{\ell} = [c_{ij}^{\ell} : i >_{\ell} j]$. The gist of the line sub-model is to relate $\mathbf{c}^{\ell}$ to $\mathbf{q}^{\ell}$, thus making up a cost-flow relationship at the line level.

*2.5. Platform issues: the passenger side*

At a platform a passenger has to wait for a route service to become available and eventually to choose a service. Let us here restrict the concept of a line to a subset of services that cannot overtake one another, especially so because they share station platforms and section tracks. Then if a passenger uses a line he will take the first service with available vehicle capacity. The line frequency on a given leg is the sum of those of the service routes that connect the leg endpoints:

$$f_{ij}^{\ell} = \sum_{z \in \ell \cap (i,j)} f_{zi} .$$

Thus service routes available along a given line are bundled together. Further combination may occur between alternative lines: assuming that if neither is saturated, then the standard condition for line attractiveness with respect to the trip destination applies. If the alternative option is a bundle $b$ with frequency $f_{i(s)}^{b}$ and minimum cost $c_{i(s)}^{b}$ up to destination $s$, then line $\ell$ is attractive if, denoting by $u_{js}$ the minimum cost from $j$ to $s$, and by $\alpha$ the discomfort coefficient of waiting time compared to in-vehicle time:

$$c_{ij}^{\ell} + u_{js} \le c_{i(s)}^{b} + \alpha / f_{i(s)}^{b} .$$

Conversely the alternative bundle is attractive if

$$c_{i(s)}^{b} \le c_{ij}^{\ell} + u_{js} + \alpha / f_{ij}^{\ell} .$$

If both options are attractive, they are combined in an extended bundle $b' \equiv b \cup \ell$ with joint frequency $f_{i(s)}^{b'} \equiv f_{i(s)}^{b} + f_{ij}^{\ell}$ and joint minimum cost $c_{i(s)}^{b'} = [f_{i(s)}^{b} c_{i(s)}^{b} + f_{ij}^{\ell}(c_{ij}^{\ell} + u_{js})] / f_{i(s)}^{b'}$.

In the unsaturated case the average waiting time is $w_{ij}^{\ell} = \alpha / f_{ij}^{\ell}$ by using $\ell$ only, or $w_{i(s)}^{b'} = \alpha / f_{i(s)}^{b'}$ if the bundle is attractive.

The saturated case occurs when there is a waiting queue that does not dissipate during the period so that a passenger is unable to access the first vehicle that arrives after his instant of joining the queue. Then the attractiveness condition for line bundling should be modified: an ad-hoc treatment is proposed in Section 4. Overall, the extended cost-flow relationship at the line level relates $(\mathbf{c}^{\ell}, \mathbf{w}^{\ell})$ to $\mathbf{q}^{\ell}$.

*2.6. Passenger traffic*

Along a network path every passenger spends time and money; he experiences state transitions e.g. between in-vehicle and pedestrian states and he is submitted to discomfort by transition as well as by state. Let us assume that all kinds of costs can be aggregated into a generalized cost which is added up along the path, including the costs of pedestrian arcs, of waiting at platform nodes and the leg costs.

Route choice along the network yields a user-optimized path of minimum generalized cost – or an optimal hyperpath if bundling occurs.

The passenger flows along the network result from the assignment to the optimal paths of the OD matrix of trip flows by origin-destination pair, $\mathbf{q} = [q_{od} : o \in O, d \in D]$ with $O$ (resp. $D$) the set of origin nodes (resp. destination).

**3. Line problems and algorithms**

The line sub-model takes a set of exogenous flows $q_{ij}^{\ell+}$ from access to egress stations as inputs to yield cost outputs $c_{ij}^{\ell}$ and $w_{ij}^{\ell}$. Internal variables include, by route service $z \in \ell$, local frequencies $f_z^{i+}$ before station $i$ and $f_z^{i-}$ after it, as well as vehicle loads $y_{sj}^{z}$ by access-egress pair $(s, j)$. It operates on the basis of a topological ordering of the stations, either from upstream to downstream or reverse from downstream to upstream.

Before stating the flow loading 'zip' algorithm (§3.3) and the leg costing 'unzip' algorithm (§3.4), let us introduce first a sub-model of transit bottleneck (§3.1) and, then, a sub-problem of platform flowing (§3.2).

### *3.1. Transit bottleneck model*

At a given station $i$ along $\ell$, consider a group of routes $Z'_i$ that serve a set of destinations $J'_i$ downstream of $i$, such that these routes and stations make up a connected component in the bipartite graph in $Z_i^\ell \times J_i^\ell$ that links the routes to the stations that they service.

Given exogenous flows $q_j^+$ by egress station, route frequencies $f'_z$ with available capacity $k'_z$ by vehicle, the objective is to yield vehicle loads $y_{ij}^z$ by vehicle servicing route $z$ and waiting times $w_{ij}^\ell$ by passenger boarding at $i$ to alight at $j$.

Define $v_j$ the stock of passengers destined to $j$ that wait on the station platform, and $n_z = \sum_{j\in z} v_j$ the number of passengers that are candidates to board on a vehicle of route $z$ when it arrives. Then the probability to board is

$$\pi_z \equiv \min\{1, \frac{k'_z}{n_z}\}. \qquad (1)$$

Assuming queuing of mingled passengers in a bottleneck, the stock variables $\mathbf{v} = [v_j : j \in Z'_j]$ satisfy a Fixed Point Problem (FPP) as follows (Leurent, 2011b):

$$\frac{2v_j}{q_j^+ H^2} = \frac{q_j^+}{v_j} - (f\pi)_j \text{, in which } (f\pi)_j \equiv \sum_{z\in j} f'_z \pi_z(\mathbf{v}). \qquad (2)$$

Some destinations may not be queued, if the exit flow $q_j^- = (f\pi)_j v_j$ matches the entry flow $q_j^+$: in such a case the FPP is only an approximation and yields $v_j << Hq_j^+$.

The FPP has a solution which is unique. It can be solved using a Newton algorithm.

From the stock variables $v_j$, the $\pi_z$ and $q_j^-$ follow, as well as the vehicle inflows $y_{ij}^z \equiv \pi_z v_j$. Furthermore, for destination $j$ the queue duration is $H_j = q_j^+ H / q_j^-$, yielding waiting cost as follows:

$$w_{ij}^\ell = \frac{v_j H_j}{q_j^+ H} = \frac{v_j}{q_j^-} = \frac{1}{(f\pi)_j}. \qquad (3)$$

Lastly, $H_j > H$ induces $(f\pi)_j < f_{ij}^\ell$, which may hold also when $H_j = H$ but some routes are saturated.

### *3.2. Platform flowing*

The flowing of both service vehicles and passengers at station $i$ along line $\ell$ involves, first, the vehicle arrivals at upstream frequency $f_z^{i+}$ and the discharge from each vehicle of a number $y_{si}^z$ of passengers from upstream station $s$ to current station; second, the revision of route frequency; third, the waiting and eventual boarding of those passengers coming in at $i$; fourth, the propagation to downstream of the refilled vehicles and the revised route frequency.

The second step deals with the constraint of platform occupancy by the dwelling and gap times of the service vehicles. The assignment period, $H$, should suffice to accommodate the boarding and alighting times of passengers plus the inter-vehicle gaps:

$$\sum_{z \text{ via } i} f_z^{i+} \omega_{zi} + \sum_{z\in i} f_z^{i+} \max\{T_{zi}^0, T_{zi}^1 + y_{zi}^- . g_{zi}^- + y_{zi}^+ . g_{zi}^+\} \le H \qquad (4)$$

Enforcing the constraint may require to reduce frequency, which is accomplished by multiplying the left hand side by a reduction factor $\eta_i \leq 1$. To simplify the evaluation of (4), let us substitute the residual capacity by vehicle after passenger discharge, $k_{zi}^{o}$, to the still unknown incoming flow $y_{zi}^{+}$. If constraint (4) is satisfied for $\eta_i = 1$ then frequency $f_z^{i-} \equiv f_z^{i+}$ is unchanged from upstream to downstream, whereas if $\eta_i < 1$ then

$$f_z^{i-} \equiv \eta_i f_z^{i+} \text{, in which} \tag{5}$$

$$\eta_i = H / [\sum_{z \text{ via } i} f_z^{i+} h_{zi} + \sum_{z \in i} f_z^{i+} \max\{T_{zi}^{0}, T_{zi}^{1} + y_{zi}^{-}.g_{zi}^{-} + k_{zi}^{o}.g_{zi}^{+}\}]. \tag{6}$$

Remark that vehicle gaps must be summed over all services that pass through that platform at station $i$, either stopping or not. All the "through" routes are submitted to frequency revision.

The third step requires solving as many sub-problems of transit bottlenecks as there are "clusters" of egress stations and routes.

The fourth and last step simply consists in propagating, for each service route $z$ passing through $i$, the revised variables $f_z^{i-}$ to the next station along $z$ ahead of $i$, together with the completed vehicle loads $(y_{sj}^{z})$ if $i$ is a dwelling station for $z$.

### *3.3. The line flow loading problem and zip algorithm*

To load both flows of passengers and vehicles onto the line is the problem of setting out the local route frequencies and the local vehicle loads of passengers by route and access-egress pair in a coherent way, by meeting the local capacity constraints at platforms and in-vehicle while giving priority to conditions upstream onto those downstream.

This problem of flow loading can be solved by a *zip algorithm* that deals with every station in turn, from upstream to downstream in the topological order of the line. At station $i$, solve the platform flowing problem on the basis of the local conditions $f_{zi}^{+}$, $[y_{sj}^{z} : s > i, j \leq i]$ inherited from all upstream stations $s$, and propagate the adapted conditions $f_{zi}^{-}$, $[y_{sj}^{z} : s \geq i, j < i]$ to the downstream stations.

The complexity in computation time of the zip algorithm is a product of the number of stations by the number of routes, times the number of operations required to solve any transit bottleneck model – the latter depends on the number of egress stations and of routes dwelling at $i$. The complexity in memory space may be limited to the number of routes times the number of egress stations, since only the passenger load by egress station and service route (not by entry station) need to be kept in memory.

### *3.4. The leg costing problem and unzip algorithm*

To cost the line legs is the problem of evaluating the travel conditions to a passenger, by pair of access-egress stations. The travel conditions include the eventual money expenses, the time spent either in-vehicle or on platform, perhaps evaluated as generalized rather than physical time to account for local discomfort.

Concerning physical time, the vehicle times by service route and the passenger waiting times at stations stem from the line loading problem. The in-vehicle section time may be taken as exogenous, i.e. $t_{za}$, or state dependent, for instance by including a time penalty at any station at which the service frequency is reduced.

Generalized time involves physical time multiplied by discomfort coefficients – say 1 for sitting in-vehicle and about 2 for waiting on an uncongested platform or for standing in a crowded vehicle. Discomfort functions with respect to passenger density on platform or in-vehicle may be considered, eventually distinguishing between sitting and standing riders as in Leurent (2011a).

Leg costing can be addressed by an *unzip algorithm* that deals with every egress station in reverse order from downstream to upstream, after evaluation by network element of the local generalized time on the basis of the physical time and the flow density. By egress station $s$:

- Initialize variables of physical time $T_{zs}$ and generalized time $G_{zs}$ to zero for all routes $z$ that pass through $s$ or to infinity otherwise.
- Enumerate the line stations $i$ upstream of $s$ in reverse topological order from $s$, (i) to cumulate $T_{zs}$ and $G_{zs}$ from $i$, (ii) to evaluate the share of passenger flow incoming at $i$ and destined to $s$ between the available routes, and the leg cost, on the basis of the following formulas:

$$p^{\ell}_{z/i,s} = \frac{f^{+}_{zi}\pi^{\ell}_{zi}}{(f\pi)^{\ell}_{i}} \text{ in which } (f\pi)^{\ell}_{i} \equiv \sum_{z\in(i,s)} f^{+}_{zi}\pi^{\ell}_{zi}\,. \tag{7}$$

$$c^{\ell}_{i,s} = \frac{1}{(f\pi)^{\ell}_{i}} \sum_{z\in(i,s)} f^{+}_{zi}\pi^{\ell}_{zi} G^{\leq i}_{zs}\,. \tag{8}$$

The computational complexity both in time and space amounts to the squared number of stations times the number of routes.

### 3.5. On the static assumption and the continuity of flow

The static framework is convenient to deal with the constraint of platform occupancy by the passage and dwelling of vehicles. However the line model is likely to disrupt the conventional assumptions in static assignment in two ways: first, the transit bottleneck model may turn $H$ into $H^{\ell}_{ij}$ for access-egress pair $(i, j)$, yielding period flow of $Hq^{(\ell)-}_{ij}$ equal to $q^{(\ell)+}_{ij} H^2 / H^{(\ell)}_{ij}$ instead of $Hq^{(\ell)+}_{ij}$; second, the local frequency revision may decrease a frequency of $f^{-}_{zi}$ at leg entry down to $f^{+}_{zj}$ at leg exit, which provides some ground for an additional delay of $\frac{1}{2} H(1 - f^{+}_{zj} / f^{-}_{zi})$ by vehicle that is operated during the reference period but also disrupts the conservation of the service flow during the reference period $H$.

As the main purpose of the line model is to establish the leg costs, these disruptions are considered by the authors to be significant mainly as results internal to the line model – which may be of separate interest to a study analyst – whereas in the context of network assignment the travel conditions should be applied to leg flows $q^{(\ell)}_{ij}$ taken only at their line entry value, $q^{(\ell)+}_{ij}$.

## 4. Network model

Let us turn to the general problem of network assignment, with passenger route choice from origin to destination pair of nodes.

### 4.1. Traffic state and hyperpath representation

Hereafter the service network with private and leg arcs is considered: the arc set is $A_S$. A *Traffic state* is a vector of arc flows, $\mathbf{x}_A = [x_a : a \in A_S]$.

The *arc travel cost function*: $\mathbf{x}_A \mapsto \mathbf{c}_A = \mathbf{C}_A(\mathbf{x}_A)$ yields, on pedestrian arcs, the local cost with respect to local passenger flow $x_a$ and, on leg arcs, the leg costs that stem from the line assignment of the leg flows in $\mathbf{x}_A$ (by transit line). This function can be made continuous by enforcing on each route and at each station a strictly positive though arbitrarily small minimal residual capacity. This is both innocuous and realistic enough since any passenger would be ready to board even in a very crowded vehicle rather than to wait indefinitely long.

If several lines are available at a given node $n$, each leg arc $a$ with post-boarding cost $u_a$ and platform waiting time $w_a$ and ex-ante frequency $f_a$ (Cf. frequency $f_{ij}^{\ell}$ in Section 2), then line combination may occur in a revised way of the uncongested model: let $\beta_a \equiv \varphi(w_a - 1/f_a)$ with $\varphi$ a continuous function decreasing from $\varphi(0)=1$ down to $\varphi(x)=0$ beyond some small argument $\varepsilon$. Let us assume that line $a$ delivers a minimum time of $u_a + \alpha(1-\beta_a)w_a$ and a mean time of $u_a + \alpha[(1-\beta_a)w_a + \beta_a / f_a]$. Then the condition for line $b$ to be attractive relative to line $a$ would be

$$u_b + \alpha(1-\beta_b)w_b \le u_a + \alpha[(1-\beta_a)w_a + \beta_a / f_a]. \tag{9}$$

Line bundling would occur between attractive lines similarly to the uncongested model but for the revised frequency, $\hat{f}_a \equiv f_a / \beta_a$, yielding:

- Combined frequency of attractive bundle $B$: $\hat{f}_B \equiv \sum_{a\in B} \hat{f}_a$.
- Passenger flow share by line of $\hat{f}_a / \hat{f}_B$ if attractive or zero otherwise.
- Bundle average cost of $u_B \equiv [\alpha + \sum_{a\in B} \hat{f}_a (u_a + \alpha\beta_a w_a)] / \hat{f}_B$.

Let us define a *hyperpath* $h \subset A$ as an oriented, connected, acyclic sub-graph that directs any of its nodes to a given destination node say $s$. Based on arc costs and leg waits and nominal frequencies, by applying the bundling treatment in recursive order from $s$, routing proportions $\gamma_a^h$ are associated to the arcs $a \in h$ that are bundled, whereas unbundled arcs in $h$ or arcs out of $h$ have routing proportion $\gamma_a^h \equiv 0$. The *hyperpath cost* of $h$ from $n$ to $s$, denoted $c_{ns}(h, \mathbf{x}_A)$, is the cost of the bundle at $n$ based on the $\mathbf{x}_A$ traffic state. Denote by $\eta(n,s)$ the set of hyperpaths from $n$ to $s$.

A hyperpath of shortest cost can be constructed recursively from $s$ by searching at each node for the optimal bundle.

### *4.2. User optimization and traffic equilibrium*

A traffic state $\mathbf{x}_A$ induces arc costs, leg waits and nominal frequencies, thus shortest hyperpath $h^*$ and associated costs $c_{ns}(\mathbf{x}_A)$ between any pair $(n,s)$ of nodes. Let $\mu_{NS} \equiv [c_{ns}(h^*, \mathbf{x}_A) : n \in N, s \in D]$ denote the vector of shortest costs.

A *Hyperpath flow vector* $\zeta$ has one component $\zeta_{ns}^h$ by hyperpath $h$ from $n$ to $s$.

It is *admissible* if $\zeta_{ns}^h \ge 0 \;\; \forall n,s,h$ and $\sum_{h\in(n,s)} \zeta_{ns}^h = q_{ns}$ the OD flow.

It is *user-optimized* with respect to traffic state $\mathbf{x}_A$ if and only if:

$$\forall n,s,\; \forall h \in \eta(n,s):\;\; \zeta_{ns}^h > 0 \Rightarrow c_{ns}(h, \mathbf{x}_A) = \mu_{ns}. \tag{10}$$

*Definition*. A *Traffic equilibrium* associated to OD demand $\mathbf{q} = [q_{ns} : n \in N, s \in D]$ is a pair $(\mathbf{x}_A, \zeta)$ such that (i) $\zeta$ is admissible with respect to $\mathbf{q}$, (ii) $\zeta$ is user-optimized with respect to $\mathbf{x}_A$ and (iii) $\mathbf{x}_A$ results from the network assignment of $\zeta$ based on $\mathbf{x}_A$, i.e.

$$\forall a \in A:\;\; x_A = \sum_{s\in D} \sum_{n\in N} \sum_{h\in\eta(n,s)} \zeta_{ns}^h \sum_{r\in a\cap h} \gamma_r^h(\mathbf{x}_A), \tag{11}$$

Wherein $\gamma_r^h \equiv \prod_{a\in r} \gamma_a^h$ is the routing proportion of path $r$ along hyperpath $h$.

This definition amounts to a fixed point problem. It is equivalent to a quasi variational inequality problem in the space of hyperpath flow vectors. As the cost function is continuous and the space has finite dimension, there must exist an equilibrium state.

### *4.3. MSA Algorithm and convergence criterion*

A Method of Successive Averages (MSA) can be used to solve the traffic equilibrium problem of network assignment in the following way, using a decreasing sequence of positive numbers $(\lambda_k)_{k\geq 0}$ with $\lambda_0 = 1$:

- *Initialization*: set $\mathbf{x}_A := 0$ and $k := 0$.
- *Cost evaluation*: by network element based on $\mathbf{x}_A$. This involves notably to deal with each line by applying the zip algorithm to load the leg flows and then the unzip algorithm to cost the legs.
- *Search for shortest hyperpaths* for all origin-destination pairs and *load the OD flows along them* (and with respect to $\mathbf{x}_A$) onto the network elements, yielding auxiliary traffic state $\mathbf{y}_A$. By destination, a shortest hyperpath is built on the service network from each node recursively, in the classical way save for the adaptation of bundle cost as in Section 4.1.
- *Next state*. Let $\mathbf{x}'_A := \lambda_k \mathbf{y}_A + (1-\lambda_k)\mathbf{x}_A$.
- Evaluate a *convergence criterion* between $\mathbf{x}'_A$ and $\mathbf{x}_A$. If it is sufficiently small then Stop with solution $\mathbf{x}'_A$, else increment $k$, replace $\mathbf{x}_A$ by $\mathbf{x}'_A$ and go to Cost Evaluation.

A convenient yet crude convergence criterion may be $\|\mathbf{x}'_A - \mathbf{x}_A\|$. A Matlab implementation is available on request from the authors.

## 5. Application instance

A classroom instance was built to demonstrate the application of the model. The case mimics the busiest railway line in the Paris metropolitan area, named RER A – the RER being the Regional Express Network. At the morning peak the directional passenger flow through the central trunk amounts to about 50,000 persons per hour; the nominal frequency of 30 trains per hour is frequently decreased to 25 on average, due to congestion which is particularly acute at central stations where many transfers take place.

In the East to West direction, there are two main service routes which link the two Northern (resp. Southern) branches to the central trunk (Fig. 1). At the morning peak hour the operation frequency is of 18 /h and 12/h respectively, yielding a nominal frequency of 30 trains per hour on the central trunk. Between North-East and Centre another railway line, RER E, competes with RER A. The passenger capacity by train is about 2,000 depending on the route. The minimum dwelling time is planned as 40 s on RER A and 50 s on RER E which is less congested.

Six stations were selected as origin or destination zones. The model of infrastructure network includes 28 nodes and 35 arcs (Fig. 2). On this small network convergence was achieved in fifty iterations (Fig. 3).

Here are some selected assignment results:

- At the Nation station, vehicle dwell time amounts to 54s (resp. 58s) on the North route (resp. South). These are reduced to 43s (resp. 42s) at Auber station where there is less residual capacity.
- From Nation station westwards, the operation frequency is reduced to 17.1/h (resp. 11.4/h) on the North (resp. South) route: this is a 5% reduction on the nominal capacity of the line.
- Individual wait time is increased from 3.3’ (resp. 5.0’) on the North (resp. South) route in the absence of congestion, to 5.6’ (resp. 6.7’) at Nation station and 11.6’ (resp. 17.0’) at Auber station, owing to the interplay of incoming flow and residual capacity.
- On the OD pair from Nation to Saint Germain (South West), at equilibrium only the South route is used, not the Northern one because the wait time at Auber is dissuasive to transfer there. This induces a significant change in the share of flow incoming at Nation between the North and South routes westwards, from 2 versus one thirds without congestion to 57%-43% under congestion.
- The travel cost from Nation to Auber is increased from 12.0’ without congestion to 14.0’ with congestion. From Auber to Cergy (resp. Saint Germain) it is increased from 43.5’ (resp. 30’) to 50.8’ (resp. 42’). The variation in generalized time would be even larger by penalizing the time spent in-vehicle in order to capture the discomfort of vehicle crowding.

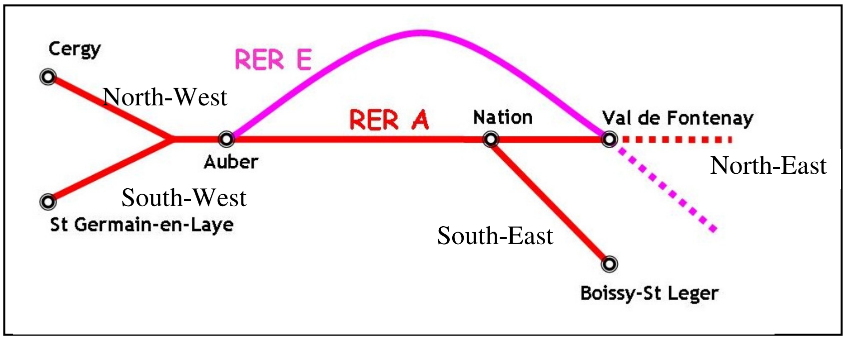


Fig. 1. Abstraction of selected network, emphasizing line legs.

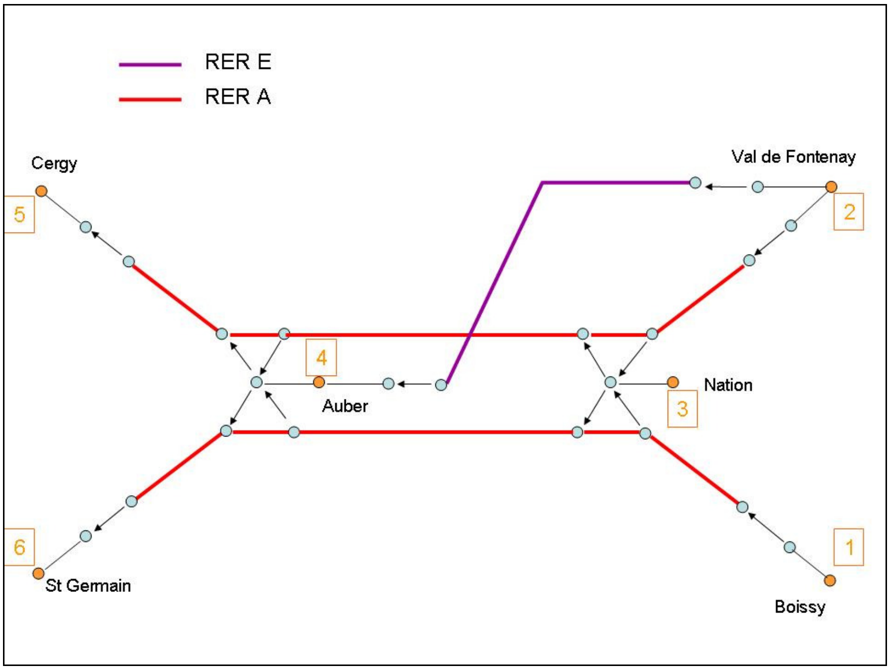


Fig. 2. Underlying infrastructure network.

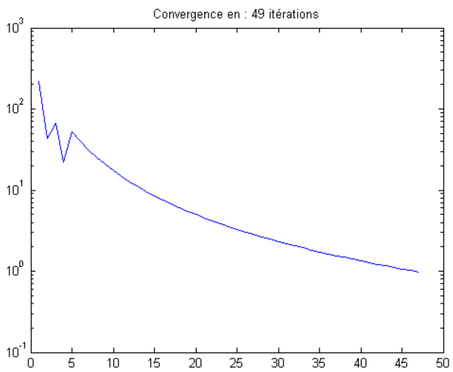


Fig. 3. Convergence criterion vs iteration number.

## 6. Conclusion

A transit assignment model has been provided, which is structural because it deals with a hierarchy of layers or scales – from passenger at the bottom to network at the top passing by vehicle, service route and line – in an explicit way, along with the spatial structure of routes and lines. Spatial features are considered at each scale, with emphasis on capacitated resources of vehicle platform as well as access-egress of passengers to vehicles. In spite of the static assumption, time features are involved in resource occupancy, frequency reduction and the associated delay, physical and generalized passenger travel times. There are model variables standing for passenger stocks on platform and in-vehicle.

The model treatment involves a series of innovations: first, the transit bottleneck model for passenger flowing at platform to vehicles; second, the influence of passenger flows on dwelling times; third, the local revision of service frequency based on platform occupancy and its downstream propagation; fourth, a revised condition for attractiveness for line combination at choice nodes; fifth, a hyperpath framework with state-dependent bundling.

The model addresses a wide range of traffic phenomena: in-vehicle passenger capacity, vehicle access-egress capacity, platform capacity on the vehicle side. It would be straightforward to include seat capacity by vehicle, corridor pedestrian capacity, track capacity in vehicles. These features are expected to be useful in transit planning applications. On-going work is focused on congestion assessment on the Paris network.

Research topics include: the estimation of passenger behaviour under congestion; the refinement of operating policy – beyond proportionate frequency reduction; the stochasticity in passenger stocks and flows, dwelling times and section run times.

***Acknowledgement**. This work is supported by a Research and Education Chair operated by Ecole des Ponts ParisTech in partnership with the Transport Organizing Authority in the Paris region (STIF). The Chair is targeted to the socio-economic analysis and modeling of passenger transport in the urban setting.*

## 7. References

Burdett, Kozan (2006), *Techniques for absolute capacity determination in railways*, Transportation Research Part B, Vol.40, pp 616-632.

Cepeda M, Cominetti R, Florian M (2006) A frequency-based assignment model for congested transit networks with strict capacity constraints: characterization and computation of equilibria, *Transportation Research Part B*, Vol. 40, pp. 437–459.

De Cea J, Fernández JE (1993) Transit Assignment for Congested Public Transport Systems: An Equilibrium Model. *Tptn. Sc*., 27(2), 133-147.

Hoogendoorn SP, Daamen W (2005). Applying microscopic pedestrian flow simulation to railway design evaluation in Lisbon. *Transportation research record*, 1878: 83-94.

Lam WHK, Gao ZY, Chan KS, Yang H (1999) A stochastic user equilibrium assignment model for congested transit networks, *Transportation Research Part B*, Vol. 33, pp. 351-368.

Leurent F (2010) On Seat Capacity in Traffic Assignment to a Transit Network. *Journal of Advanced Transportation*, 2010;9999:1-27. (in press)

Leurent F (2011a) On capacity and congestion in passenger transit systems: Systems analysis and implications for network assignment models. *European Transport Research Review* (in press)

Leurent (2011b) *The transit bottleneck model*. Working document, Ecole Nationale des Ponts et Chaussées, Université Paris Est.

Lai, Wang, Jong (2011), *Development of Analytical Capacity Models for Commuter Rail Operations with Advanced Signaling Systems*, Paper #11-3424 presented at 90th Annual Transportation Research Board Meeting, Washington DC, January 2011.

Ortuzar JD, Willumsen L (2004). *Modelling Transport*. Wiley, 3rd edition.

Spiess H, Florian M (1989) Optimal strategies: a new assignment model for transit networks. *Transportation Research B* 23: 83-102.

Sumalee A, Tan Z, Lam WHK (2009) Dynamic stochastic transit assignment with explicit seat allocation model, *Transportation Research Part B* 43: 895–912.

Thomas R (1991) *Traffic Assignment Techniques*. Avebury Technical, Aldershot, England.

TRB (2003) *Transit Capacity and Quality of Service Manual*. On-line report prepared for the Transit Cooperative Research Program, available on-line at the following website address: http://gulliver.trb.org/publications/tcrp/tcrp_webdoc_6-a.pdf . First edition 1999.

Vuchic VR (2006) *Urban Transit : Operations, Planning and Economics*. Wiley.